DOCUMENT RESUME

ED 297 570 FL 017 212

AUTHOR Ingram, D. E.; Clapham, Caroline
TITLE ELTS Revision Project: A New International Test of English Proficiency for Overseas Students.
INSTITUTION Lancaster Univ. (England). Inst. for English Lanugage Education.
SPONS AGENCY British Council, London (England).
PUB DATE Jan 88
NOTE 31p.; Paper presented at the combined Annual World Congress on Language Learning of the Federation Internationale des Professeurs de Langues Vivantes (16th) and Biennial National Languages Conference of the Australian Federation of Modern Language Teachers Associations (7th, Australian National University, Canberra, Australia, January 4-8, 1988).
PUB TYPE Reports - Descriptive (141) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *English (Second Language); English for Academic Purposes; Foreign Countries; Foreign Students; Higher Education; International Cooperation; *Language Proficiency; *Language Tests; *Standardized Tests; *Test Construction

ABSTRACT

The English Language Testing Service (ELTS) Revision Project is a joint British and Australian effort to develop a new test of English language proficiency for use with foreign students. The existing test, in use for seven years, is being radically revised to combine indirect, semi-direct, and direct test types and to focus on both general and academic language proficiency. The report explains the project's purpose, discusses Australia's interest in the test's revision, describes the current test's construction and use, and outlines the revision project. The latter section discusses the reasons for making substantial changes, the pattern, content, and administration of the revised test, the research team, and the future of the project. The new test is scheduled for initial use in Britain, Australia, and Canada in 1989. It is concluded that the revised test will be an essentially new test reflecting the state-of-the-art in language testing. It will be administered to tens of thousands of foreign students each year and may substantially affect students' courses of study. It is also considered to be commercially viable and subject to ongoing research because of the data base it will generate. (MSE)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

ED297570

# ELTS REVISION PROJECT: A NEW INTERNATIONAL TEST OF ENGLISH PROFICIENCY FOR OVERSEAS STUDENTS

D. E. INGRAM

Director, Institute of Applied Linguistics,
Brisbane College of Advanced Education

and

CAROLINE CLAPHAM

Research Associate, University of Cambridge
Local Examinations Syndicate, ELTS Revision Project,
University of Lancaster

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

D. Ingram

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Paper to the 16th FIPLV World Congress on Language Learning/7th Biennial AFMLTA National Languages Conference, "Learning Languages is Learning to Live Together", Australian National University, Canberra, 4-8 January, 1988.

Dr. Ingram is President of the Australian Federation of Modern Language Teachers Associations and a Vice-President of the Federation Internationale des Professeurs de Langues Vivantes. From October 1987 to October 1988, he is the Australian representative on the ELTS Revision Project located at the University of Lancaster, England. Australia's participation in the project is funded by the International Development Program of Australian Universities and Colleges, Churchill House, 218 Northbourne Avenue, Braddon, A.C.T., 2601.
Caroline Clapham is a Research Associate appointed by the University of Cambridge Local Examinations Syndicate as the chief full-time British researcher on the ELTS Revision Project. The project is directed by Dr.J. Charles Alderson, is funded by the British Council, and is located in the Institute for English Language Education at the University of Lancaster, England.

FL017212

# THE ELTS REVISION PROJECT: A NEW INTERNATIONAL TEST OF ENGLISH PROFICIENCY FOR OVERSEAS STUDENTS

D. E. INGRAM and CAROLINE CLAPHAM

## ABSTRACT

In the ELTS Revision Project, Britain and Australia are cooperating to develop a new test of English language proficiency for use with overseas students. The authors are the two full-time British and Australian researchers on the team directing the project. The ELTS Test has been in use for some seven years, has been extensively reviewed, and is now being radically revised. The new test combines indirect, semi-direct and direct test types and focusses on both general proficiency and English for Academic Purposes. After extensive trials in 1988, the test is expected to come into use with overseas students coming to study in Britain and Australia from 1989.

CONTENTS

I. Purpose of the Project

II. Australia's Interest

III. The Current ELTS

IV. The ELTS Revision Project

IV.1. Reasons for Change

IV.2. Pattern of the Revised Test

IV.3. Contents of the Revised Test

IV 4. Administration of the Revised Test

IV.5. The Research Team

IV.6. The Future of the Project

V. Conclusion

Appendix One : Overall Band Descriptors

Appendix Two : Course Categories and Recommended Band Scores

Appendix Three : Target Groups and ELTS Components

Appendix Four : Administration and Marking of the revised ELTS

References.

I. Purpose of the Project: The recent rapid increase in the numbers of overseas students coming to Australia and the continuing large numbers going to Britain have as their corollary the need for more satisfactory (i.e. more valid, reliable, efficient, and informative) tests of the potential students' English proficiency. The need is three-fold: first, to assess whether students entering training programmes of different sorts have sufficient English proficiency to participate successfully in their programmes without their English skills militating unduly against their success; second, to provide a basis on which students can be advised on the nature and length of the English courses they will require to reach a proficiency level acceptable for entry to their training programmes; and, third, in the Australian context (though less so in Britain because of the availability of other tests), there is need for a test of general proficiency which will give ELICOS centres (i.e., public or private schools providing English Language Intensive Courses for Overseas Students) advance notice of the likely level (and perhaps nature) of courses suitable for the large numbers of individuals applying to come to Australia to learn English. The magnitude of the need is reflected in the following facts: first, in 1985-6, 13,121 students (mostly going to Britain) took the ELTS test, a number that had almost quadrupled since 1980-1; second, a recent report from the University Grants Committee [reported in The Times newspaper, 14 December 1987, p.4] set the total number of overseas students attending British Universities at 42,000 with many more attending Polytechnics and Colleges of Further Education or undertaking other training programmes. Third, the Australian need has expanded

dramatically since the easing of student visa regulations in 1985 and since the Federal Government has begun encouraging public and private institutions to attract full-fee-paying students from overseas. Nevertheless, in 1985-6, it was estimated that 7,818 overseas students studied English at accredited institutions in Australia [Lo Bianco 1987: 97]; in the period January to October, 1986, approximately 7,000 students took the Australian Government's "Short Selection Test"(SST) [Burke 1987: 65]; and, even though it is not required for students coming to study English in Australia, some 10,000 students currently take the SST each year [Lo Bianco 1987: 184]. The magnitude of the need can be further illustrated by two practical examples: one of the present writers was recently informed by the Australian Vice-Consul (Education) in Hong Kong that he had just administered the SST to 1,000 students over two week-ends [Peter Blackburn, personal communication, September 1987] and, in the first part of 1987, the same writer's own newly formed ELICOS unit, only one of over 70 such centres in Australia, was receiving over 100 applications each day from overseas persons wishing to study English in Australia while many other longer established centres were receiving even larger numbers. Whether or not all such applicants eventually enter courses, as applicants, they all have to be advised on the suitability of available courses to meet their needs, needs which, in the absence of informative test results, cannot be satisfactorily identified.

In view of the large numbers of students involved, the need is for a test that can be readily and rapidly administered, preferably by clerical (rather than more highly paid professional) staff, and able to be administered in remote countries and localities by persons with minimal training and subject to

relatively little expert supervision. Since the students come from many countries, since they are scattered widely in their countries, since they wish or need to enter English and other courses at almost any time throughout the year, and since they often need the test results urgently, the test must be available at very frequent intervals or on demand [ELTS 1987: preface, 3], it must be readily available anywhere (currently the British Council offers it at 144 centres in 97 countries [ELTS 1987: 3, appendix]), and the candidate's answers must be able to be marked very rapidly and, preferably, locally. In addition, once the test has been taken and scored, its results must be readily interpretable.

The ELTS test, run jointly to date by the British Council and the University of Cambridge Local Examinations Syndicate, has been in use since 1980 and was designed to try to meet the needs just outlined. It seeks to measure a candidate's general language skills but also those skills needed for study or training [cf. ELTS 1987: preface]. In the seven years it has been in use, it has gained wide acceptability, as is indicated by the rapid growth in the numbers of candidates taking it [ELTS 1987: 2] and by the growing list of institutions not only in Britain but around the world that accept it. In Britain, for example, the ELTS information materials list almost 70 universities or university colleges and 16 polytechnics that accept it [ELTS 1987: 8; ELTS 1987c]. Nevertheless, for reasons to be considered subsequently, the decision was taken in 1986 to revise the test but to keep its broad framework and, especially, its method of reporting results.

II. Australia's Interest: Australia's interest in participating in the ELTS Revision Project originates from advice provided to the Overseas Students Office of the Commonwealth (i.e. Australian) Department of Education by Dr. Ed Burke, who had been commissioned to review world practice in assessment of the English proficiency of overseas students and to advise on "the assessment of the English language abilities of private overseas (non-native English speaking) students" wishing to study in Australia (Burke 1987: 1). He was also required to advise on how responsibility for English language assessment procedures could be transferred to an Australian agent or agencies (Burke 1987: 1-2). In a long and detailed report, Burke reviews the major tests currently in use for the prescribed purpose and concludes by recommending that the SST no longer be used and that, rather than adopt any existing test or develop its own new test independently, Australia should cooperate with the British Council in revising the ELTS test. His principal reasons for recommending this cooperative, international approach included the desirability of reducing the time and cost entailed in developing a new test independently, the international acceptance the ELTS test had already achieved, the relatively progressive nature of the ELTS test itself with its combination of direct and indirect techniques, and the prospect the revision project offered of further accommodating recent developments in testing. In addition, there was the advantage to both Britain and Australia of sharing a new international test of sufficient significance to rival TOEFL and the advantage in terms of initial cost, on-going development costs, administration, and world coverage if Britain and Australia (possibly together with Canada) shared the developmental and administrative responsibilities of the test. Burke further recommended that responsibility for

Australian participation be taken up by the International Development Program of Australian Universities and Colleges (IDP) and that IDP eventually establish a small consortium of institutions to assume responsibility for the test, for its administration, and for the on-going development and validation activities required to produce multiple future forms [see Burke 1987, especially pp. 64-71]. A further dimension is added to Burke's recommendations and makes them financially viable if one considers Lo Bianco's independently made recommendation in his National Policy on Languages [1987] that a language testing unit be established, possibly as a part of a national key centre on Applied English and his observations that such a unit should be self-funding through fees charged to students taking tests; he sugggests it should be financially viable to the point where it should recoup initial establishment costs in the first year of operation [1987: 184]. In August 1987, Burke's recommendations were considered by a meeting of twenty-one teachers and applied linguists involved in language testing or ELICOS programmes and were unanimously (though not unquestioningly) supported. IDP together with a small steering committee appointed by the meeting further developed the proposal for Australian participation in the ELTS Revision Project in discussion during the 1987 AILA Congress in Sydney with Gill Westaway of the British Council and Dr. Charles Alderson, Director of the Project. Though the final details of Australia's participation in the total project and the on-going development and administration of the test have yet to be resolved, arrangements made to date entail Australia's buying into the project for a percentage of the previous

development costs, the assignment of one Australian recognized in the area of second language testing to work on the project in Lancaster for 12 months, one or two additional short-term Australian attachments to assist on the project at critical periods and to assist in relating Australian needs to the project team, and on-going participation in development and validation activities in Australia coordinated through a small Working Group. David Ingram was appointed to the 12 months assignment and commenced duties in Lancaster on 19 October, 1987.

III. The Current ELTS: The current ELTS (i.e. the one to be replaced by the Revision Project) has two components, G and M. The G (General) Component is intended to assess general language proficiency while the M (Modular) component seeks to test language skills in particular subject areas and is, to that extent, concerned with the measurement of "special" or "academic" purpose proficiency [see ELTS 1987: 5-6 for more details than presented here]. All candidates take the G component and select the M component according to their subject area. The test takes approximately 3 hours to administer and consists of the following sub-tests with test types as shown in brackets (m.c. = multiple choice).

| G | | M | |
|---|---|---|---|
| G1 | Reading (m.c.) | M1 | Study Skills (m.c.) |
| G2 | Listening (m.c.) | M2 | Writing (2 extended writing tasks) |
| | | M3 | Interview (individual guided discussion) |

The M sub-tests exist in modules covering:

Life Sciences
Medicine
Physical Sciences
Social Sciences
Technology
General Academic (taken by candidates whose area of study is not covered by one of the other modules).

Scores are reported on a 1-9 band scale (see Appendix One) with a profile covering each of the five sub-tests and an overall band score derived from the mean of the sub-tests. There is also a Non-Academic Module for candidates whose training programme will be non-academic, practical or technical in nature. Its scores are reported on a 1-6 band scale (from "1 - unable to communicate in English" to "6 - generally effective command of the language" [ELTS 1987b: 10]. The test consists of:

M1 Listening
M2 Reading and Writing
M3 Oral Interview

The band descriptions give assistance to the end-users of the tests in interpreting the results. The fact that both a profile and an Overall Band Score are provided gives more information than a single figure result. This might be of more use to English training centres than to academic programmes since there is some evidence [e.g. see ELTS 1987c, Criper and Davies 1986, Clapham 1987: 8] that institutions generally use only the Overall Band Score for entry to academic programmes. In principle, the profile would also seem to be useful for students who require to improve their English skills since it may give an indication of the aspects of their skills they need to concentrate on. In addition to the descriptions contained in the profile and Overall Band

Score, the British Council, on the basis of its experience of placing overseas students in courses, has categorized courses into four types according to their linguistic demands and the desirable Overall Band Scores (ELTS 1987: 7). These recommendations are shown in Appendix Two.

The current ELTS has been in use since 1980 and has been subject to a number of validation studies, with one in particular being produced by Edinburgh University in 1986 [Criper and Davies 1986 referred to here as ELTSVAL] and focussing on the test's practicality, validity and reliability. The Edinburgh validation study [see Clapham 1987 for a fuller summary of the report] criticized the test for being over-long, taking three hours to administer, requiring a ten-minute oral interview for which examiners have to be trained, and requiring locally trained assessors for the direct writing test. The large number of components (two G components plus three M components each of which have six alternatives) was considered confusing and time-consuming to handle. Candidates had difficulty selecting the appropriate module; yet, if they do the wrong module, then the advantage of having a "specific purpose" test is lost. Face validity was good, with both candidates and their subject supervisors in the receiving institutions generally feeling that the test was appropriate. Construct validity was considered adequate insofar as ELTS is an ESP test based on needs analysis but it was considered to fall down, first, in the same way as the theory of ESP falls down (e.g., when one attempts to identify significant discrete entities in the different registers) and, second, because no empirical needs analysis was apparently used as the basis of test construction. This in turn implies deficiencies in content validity but, curiously, the specifications on which the test was based are no longer

available (if they ever were recorded). Concurrent validity was high (around .8) with the English Proficiency Test Battery and English Language Battery but "inconclusive" between candidates' scores and the ratings assigned by English tutors and subject supervisors. The ELTSVAL researchers considered predictive validity to be adequate though the correlation with end-of-course results was only .34. Reliability was not discussed in detail but three studies showed that inter-marker reliability in M (Writing) was low (around .6), test-retest reliability in one small study was not significant and in a larger one was higher (around .5) but still low and varied between subject areas.

Following the ELTSVAL study, a conference of language testers reacted to the report [proceedings are shortly to be published but a report is available in Clapham 1987: 6-7] and, in summary, observed that the test needed to be shortened and its administration simplified, reliability (especially in the direct components of M2, Writing and M3, Interview) needed improvement, face validity was high, washback effect was good, and the scoring system facilitiated interpretation. Construct and content validity posed problems, the former because of the problems inherent in the concept of "specific purpose proficiency" and the latter because specifications were not adequate to allow test replication.

End users of the ELTS were also consulted [see Clapham 1987: 7-11] and included receiving institutions, overseas test administrators, language testers, and pre-sessional English teachers. The general feeling was that the broad design of the test should not be changed, especially the modular structure and use of band scores. However, the test was considered costly to administer, it needs to be simplified for handling purposes (hence the number of modules

needed to be reduced), it needed to be able to be administered by untrained people, and it needed to allow for the direct components to be centrally marked (i.e., in Britain or Australia). Examination of the subject intentions of the students revealed a wide range of subjects with numerous branches within many of them (e.g., 34 branches of Engineering). This suggests an almost insuperable problem for tests attempting to implement an ESP theory with the probability that any test devised for a broad subject area such as Engineering would be inappropriate for many candidates who would take it.

## IV. ELTS Revision Project:

IV1. Reasons for Change: As a result of the issues just discussed and, not least, recognition that there had been changes in attitudes to testing (not least to the suitability of multiple choice item types), to needs analysis, and, especially to the strength of ESP theory, it was resolved to revise ELTS but to maintain its overall structure and means of reporting results in a band score profile. There had been particular agreement on the need to include a G component and, in fact, Australian needs, especially in the ELICOS area, probably require further strengthening of this component. Face validity probably requires inclusion of modular components though they also create false expectations in candidates who think they should be specific to their own specialist fields, an impractically large requirement given the diversity of subject areas and their specialisations. In addition, there is little convincing evidence available to support the ESP testing hypotheses that different disciplines and sub-disciplines require distinct linguistic skills and, therefore, specific tests,

that it is possible to devise such specific tests, or that students are disadvantaged by taking a test not specifically in their subject area. In addition, a specific purpose test used for entry purposes, if based on the linguistic skills used in the discipline to be entered, would seem to be assessing skills yet to be experienced and developed whereas an entry test would seem most appropriate if it assessed the extent to which non-native speakers were able to perform comparably to native speakers entering the discipline. This reinforces the importance of the G component and the questionnable value of attempting to make the modular components too specific, even though face validity may favour the use of some modules purporting to assess "specific purpose" proficiency. Furthermore, studies [reported in Clapham 1987: 11-14] have yielded, at best, inconsistent results when the effect of academic background on different test performance was considered though Alderson and Urquhart [1985] did find that certain subject areas tended to cluster, e.g., business studies and humanities, engineering and science.

IV.2. Pattern of the New Test: For the reasons just discussed, the decision was taken to keep the G and M structure, to strengthen the G and, taking account of both the nature of the discipline and the number and distribution of candidates, to reduce the M to four modules, Arts and Social Science, Physical Sciences and Technology, Life and Medical Sciences and a so-called "non-academic" module.

One effect of reducing the number of modules is to make them less specific and, therefore, to make them more acccessible to both graduate and undergraduate students whose academic skills requirements are similar. These candidates would take both the G and the most relevant M components. It was

felt that the M components with their emphasis on academic skills in broad discipline areas were not suitable for students entering vocational training with emphasis on apprenticeships and on-the-job training or for persons participating in attachments. These people will take the G component as well as a "non-academic" module in which reading and writing skills will be assessed. Sub-tertiary students, those entering "access" programmes leading to later entry into academic courses, were considered unlikely to be able to cope with the M components and they will take G only. On the other hand, it was thought that students from countries where English is the medium of instruction and a major second language in the community (e.g., Papua New Guinea and many Commonwealth countries) are not expected to take G, but, because many have difficulties with study skills, would take a relevant M test. In the Australian context, it is hoped that the revised ELTS test may be available for use with students entering ELICOS programmes though some broadening of the G component will be needed if this is to occur. If it does occur, then students entering solely to take a general English course would take G while those entering to take an EAP or ESP course probably leading to subsequent entry to an academic course would take both G and M. This pattern is shown in Appendix 3.

IV.3. Contents of the Revised Test: It should be noted that, at the time of writing, draft specifications and draft item types for inclusion in the

revised test are just being drawn up, have not yet been approved by the Steering Committee, and may be changed.

The revised test contains the following sub-tests:

| | |
|---|---|
| G1. Lexis and Structure | M1. Academic Reading |
| G2. Listening | |
| G3. Oral Interaction | M2. Academic Writing |

Non-Academic

M1 and M2 are realized in three discipline areas as indicated above: Arts and Social Sciences, Physical Sciences and Technology, and Life and Medical Sciences. Despite their possible advantage in objectivity of marking even by non-professionals, the use of multiple choice item types is to be minimized.

Reference has already been made to the importance of and constraints on, needs-analysis in devising ESP tests. The approach in the Revision Project is to use teams of persons with known expertise and experience in language testing, in pre-sessional courses, or in the discipline areas. Draft specifications and draft items are being prepared on the basis of their knowledge and experience and these are subsequently to be sent to language testers, subject specialists, receiving institutions, test administrators, and anyone else relevant for feedback and advice on the extent to which what is proposed meets the needs.

The inclusion of G1 (Lexis and Structure) as a sub-test partly reflects the concerted views of the informants that such a test should be incorporated

and partly is an acknowledement that, despite the recent emphasis on functional language, all language events are realized through lexis and structure and there is a place in proficiency assessment for such a test. There is also a certain amount of evidence that inclusion of such a test in a proficiency battery enhances validity and reliability (see Clapham 1987, 18-19). The current proposals (yet to be subjected to preliminary trials or finally approved) envisage a test in three components, one focussing on "Structure in Context", one on Lexis, and one on "Reference, Cohesion and Discourse Features". In order to ensure that lexis and structure are assessed in context, gap-filling types (essentially selective and random cloze) are being proposed.

Many informants felt that the G2 (Listening) sub-test should be included in M because of the importance for students of being able to listen to a lecture, take notes, and carry out a writing task. However, it is not practical, in test administration terms, to have different listening tests for the diverse array of candidates in each test centre. Consequently, Listening is in the G component. Current proposals envisage a test in three sections, one focussing on " social listening", one on "general listening" relevant to students, and one on listening in an instructional or study context. A variety of response formats are to be used including information transfer (such as form-filling, completing a diagram, following routes on a map), more traditional true-false multiple choice types, and open-ended answers to questions. It is expected that listening texts will be recorded in accents relevant to the candidate's target country (e.g., Britain, Australia or, possibly, Canada).

The G3 (Oral Interaction) sub-test is to be an oral interview conducted by trained interviewers, recorded, and assesssed either locally or in the target

country (i.e., Britain, Canada or Australia). This test is being devised as a direct test of oral proficiency but with more constraint on the interviewer in order to try to improve validity and reliability. The interview lasts ten to fifteen minutes, it will be in five phases, and it includes activities that progressively extend the candidate and give him or her the opportunity to lead the discussion. After a short introductory phase in which the interviewer elicits basic personal information, Phase One uses cue cards to have the candidate elicit information about the target country and solve a problem posed; Phase Two uses a visual stimulus to elicit extended speech; and Phase Three draws on a short curriculum vitae the candidate supplies the interviewer and which is used to elicit "speculative" information, attitudes and intentions. There is a very short concluding phase entailing little more than good wishes and farewell. Assessment is by matching observed language behaviour against a band scale containing nine brief performance descriptions from 1 (Non-Speaker) to 9 (Expert Speaker) but, in order to increase reliability of rating, as each phase is concluded, the interviewer rates it against a sub-scale focussing on different features in each phase ("Initiative and Appropriateness" in Phase 1, "Accuracy and Fluency" in Phase 2, and "Comprehensibility and Lexical Range" in Phase 3). The final rating may be the mean of the sub-scales but it is emphasized that, in the event of mismatch, the overall rating for the whole interview using the comprehensive band scale takes precedence. This sub-test holds promise of being particularly interesting from a test design point of view, drawing on the developments in "direct", interview-based assessment of the last decade but seeking to control the interview to maximize validity and

reliability for administration on a mass scale in remote locations using minimally trained interviewers.

At the time of writing, the modular sub-tests are not so advanced as the G component and the non-academic module has had to await work on the G component. The academic modules assess reading and writing in an ESP/EAP context. In each case, the responsible test development team has sought to resolve the problem of relevant text selection and to identify the use made of language in the broad discipline areas covered by each sub-test. Thus, it is expected that each sub-test might include activities related to, for example, note-making, argumentation, skimming and scanning, or information transfer (e.g., from graph to text).

The proficiency level of the revised test is a critical issue determining the extent of its usefulness, especially for Australian requirements. The briefing notes to the teams of test writers indicate that the M tests should "discriminate best at Band Level 6.0" (about 3 on the Australian Second Language Proficiency Ratings) and that the G test should "discriminate best at Band Level 4.0" (about 1+ on the ASLPR). Nevertheless, the test is not seen as a simple threshold test to indicate acceptability or unacceptability for entry to courses since it is necessary to be able to advise unsuccessful candidates on the length and type of pre-sessional English course they are likely to need. This requirement is similar to that of ELICOS Centres which want to know applicants' proficiency levels and likely course needs before their arrival. For these purposes, the revised ELTS must assess across a wider proficiency range than just Bands 4 or 6. In fact, all teams working on G sub-tests are aware of this need, it is certain that G3 (Oral Interaction) will meet the need, but it

is not yet clear to what extent the other draft specifications and items will do so. A related issue relates specifically to reading and writing. If, for the moment, one puts aside the question of test overlap, especially with G1 (Lexis and Structure) which uses cloze-type procedures, reading and writing are assessed in the M component which assumes a higher proficiency level and assesses that proficiency in EAP/ESP contexts. Thus the current proposals for the revised test do not envisage a measure of proficiency being given in reading and writing at ASLPR levels 0+ to 2+. This problem is being re-considered at the time of writing especially in the context of ensuring that the revised ELTS meets Australian (especially ELICOS) needs more comprehensively.

IV.4. Administration of the Revised Test: The need, as noted earlier, is for a test that can be administered to large numbers of students, in many countries, often in remote localities, and often by administrators who may be native speakers but who, in many cases, will have had no language teacher training. Accordingly the direct tests, writing and speaking, may be centrally marked in the target country (Britain, Australia or Canada) and some training using a self-directed training package is to be prepared for all administrators, especially those responsible for interviewing for G3. In addition, to enable the tests to be marked locally and at minimum cost, all sub-tests except G3 and M2 are being designed to be marked by clerical staff and, hence, there has to be reasonable objectivity in the marking required or any judgements involved must make reasonable expectations of untrained native speakers. These administration and marking constraints are shown in Appendix Four.

IV.5. The Research Team: On the British side, the ELTS Revision Project is administered by the British Council with the cooperation of the University of

Cambridge Local Examinations Syndicate. In Australia, the project is managed by the International Development Program of Australian Universities and Colleges (IDP) though Brisbane College of Advanced Education has also provided significant assistance with finance and release of personnel. In addition to the allocation of personnel referred to earlier, Australian involvement entails setting-up an Australian-based Working Party to receive and respond to information on the project and to manage development activities in Australia (e.g., distribution and gathering of responses to the draft specifications, preparation of additional items for use in the trial tests, and the Australian components of the validation studies). The teams principally responsible for the Revision Project are located in Britain. Overall responsibility is in the hands of a Steering Committee consisting of representatives of the British Council, the University of Cambridge Local Examinations Syndicate, the Australian representative on the research team, and both the Director and Research Associate from the Revision Project team. The key research team is located at the University of Lancaster and consists of Dr. J. Charles Alderson, Project Director, Caroline Clapham (Research Associate), and Dr. David E. Ingram (Australian representative). It is expected that a Canadian tester, Stan Jones, will join this team at intervals throughout the project. A further short-term Australian attachment (yet to be named) will join the team during the crucial phase in April-May 1988 when the final specifications are being resolved. In addition, for each G and M sub-test, a team of three people selected from around Britain has been established to draw up test specifications, draft items, and, eventually, the trial tests. The meetings of these teams are also attended by an observer from the Steering Committee and by the Australian representative

(who is also a member of the G3 team). Finally, a team of three "moderators" has been established to review the draft specifications for the M sub-tests in order to ensure "that some link is maintained between the 3 academic M's" (see Briefing Notes for the M teams, p.2).

IV.6. The Future of the Project: A number of issues remain to be resolved in the Test design. These include the nature of the "non-academic" module, whether G can be strengthened to meet Australian ELICOS needs (especially by broadening the proficiency range assessed and including reading and writing), whether the revised test can be an international test or whether cultural factors and differing national needs require separate national tests, a name for the revised test to reflect its international origin and status, ownership of the test in view of the participation of two and possibly three nations in the project, and how long-term administration of the test and on-going development of parallel forms are to be managed and costed. In Britain, the British Council is ultimately responsible for the test's management and administration but, in Australia, though IDP has accepted responsibility for the development phase, it is probable that the longer term management, administration and development of the test will be assumed by a small consortium of institutions possibly including IDP.

To date, the ELTS Revision Project has assessed the validation studies on the current ELTS, considered the views of relevant persons and institutions on the directions the Revision Project should take, and established the teams who

are developing draft test specifications and draft items. When these are complete, they will be sent to relevant people (receiving institutions, subject specialists, testing experts, and others) to elicit feedback and advice on their appropriateness. The final specifications will then be prepared and teams of test writers in Britain and, it is hoped, in Australia will prepare items for the trial tests. The validation study will be conducted in the last months of 1988 and final specifications and tests will be prepared early in 1989 by which time training manuals for test administrators and markers will also be developed. Thereafter the test will be ready for administration and, if not before, the structures to enable its on-going management, administration and development of parallel forms will need to be established.

V. Conclusion: The test that emerges from the ELTS Revision Project will, except in broad structure and method of reporting results, be essentially a new test. When it appears in 1989, it will be expected to reflect the "state-of-the-art" in language testing (though the collaborative, team-based structure of the project is more likely to produce a "solid" and conservative rather than highly innovative test). As a project, it is already attracting interest from around the world, as is to be expected from a project that brings together diverse experience and expertise, in language testing from already two and potentially four countries (in fact many more if the countries of experience of the participants are included). It is also a project of major significance, which will produce a test to be used as the principal means of assessing some tens of thousands of overseas students each year coming to study in two and probably three countries (Britain, Australia and Canada). It will, further, be a

test that will directly and profoundly influence the lives of, literally, millions of students from many countries around the world: it will directly affect those hoping to go abroad to study but, through the washback effect of the test on ESL/EFL programmes intended for students who may wish subsequently to apply for overseas studies, it will indirectly influence the courses of study of many more. Because of the large numbers of potential candidates for the test, the revised ELTS is likely to be commercially viable and, for this reason, because of the sheer magnitude of the data base it will generate, and because its on-going management and development will require on-going research, the revised ELTS will undoubtedly provide a new stimulus to research and development in second language testing in the participating countries.

## APPENDIX ONE

### Overall Band Descriptors

The results of the current ELTS are reported in the form of an Overall Band Score derived from the mean of the individual Band Scores for each sub-test. The band descriptors are as follows:

9. Expert user: fully operational command of the language; appropriate, accurate and fluent with complete understanding.

8. Very good user: fully operational command of the language; occasional minor inaccuracies, inappropriacies or misunderstandings possible in unfamiliar situations.

7. Good user: operational command of the language, occasional inaccuracies, inappropriacies and misunderstandings in some situations.

6. Competent user: generally effective command of the language, although occasional misunderstanding and lack of fluency could interfere with communication.

5. Modest user: partial command of the language coping with overall meaning in most situations although some misunderstanding and lack of fluency could block communication.

4. Limited user: basic functional competence limited to familiar situations, but frequent problems in understanding and fluency can make communication a constant effort.

3. Extremely limited user: below level of functional competence; although general meaning can be conveyed and understood in simple situations there are frequent breakdowns in communication.

2. Intermittent user: no real communication possible although single word messages may be conveyed and understood.

1. Non-user: unable to use the language or does not provide relevant evidence of language competence for assessment.

(Adapted from ELTS 1987: 6)

APPENDIX TWO

Course Categories and Recommended Band Scores

| Category | Courses | Overall Band Scores | |
|---|---|---|---|
| | | Acceptable | Probably Acceptable |
| A | Linguistically exacting academic courses, e.g. medicine, linguistics, law. | 7.5 | 7.0 |
| B | linguistically less-exacting academic courses, e.g. engineering, agriculture, business studies. | 7.0 | 6.5 |
| C | linguistically exacting training courses, e.g. air traffic control, media production. | 6.5 | 6.0 |
| D | linguistically less-exacting training courses/industrial and agricultural attachments. | 5.5 | 5.0 |

[Adapted from ELTS 1987: 7]

APPENDIX THREE

Target Groups and ELTS Components

| Target Group | G | M |
|---|---|---|
| Graduate and Undergraduate | Yes | Yes |
| Vocational Training, Apprentices, On-the-job Training | Yes | Non-academic |
| Sub-tertiary, "access" | Yes | No |
| Attachments | Yes | No or non-academic or depends on the nature of the attachment. |
| ELICOS (General English only) | Yes (if broadened) | No or possibly non-academic |
| ELICOS (EAP/ESP) | Yes | Yes |
| Students from E2L Countries | No | Yes |

APPENDIX FOUR

Administration and Marking of the Revised ELTS

| Sub-test | Administration | Marking |
|---|---|---|
| G1 (Lexis and Structure) | Clerical | Clerical |
| G2 (Listening) | Clerical | Clerical |
| G3 (Oral Interaction) | Trained ELT specialist or Trained non-ELT-specialist | Trained ELT specialist either at local centre or centrally in target country (Britain, Australia, Canada) |
| M1 (Reading) | Clerical | Clerical |
| M2 (Writing) | Clerical | Trained (ELT specialist either at local centre or centrally in target country (Britain, Australia, Canada) |

REFERENCES

Alderson, J. Charles and Caroline Clapham. 1987. "The Revision of the ELTS Test". Paper to the English Language Education Forum, University of Lancaster, 3 November, 1987.

Alderson, J. Charles and A.H. Urquhart. 1985. "This Test is Unfair: I'm not an Economist". In Hauptman et al 1985.

Burke, Ed. 1987. "Final Report on the Evaluation of Overseas Standardized English Language Proficiency Tests: Implications for English Language Proficiency Testing in Australia". Report to the Overseas Students Office, Commonwealth Department of Education, Canberra, Australia, 30 April, 1987.

Clapham, Caroline. 1987. "The Rationale for the Structure of the Revised ELTS". Mimeograph.

Criper, Clive and Alan Davies. 1986. "Edinburgh ELTS Validation Project: Final Report". Mimeograph.

English Language Testing Service (ELTS). An Introduction to ELTS. London: British Council.

English Language Testing Service (ELTS). 1987a. Specimen Materials Booklet. London: British Council.

English Language Testing Service (ELTS). 1987b. Non-Academic Training Module: User Handbook. London: British Council.

English Language Testing Service. 1987c. "Acceptability of ELTS in U.K. Universities and Polytechnics." London: British Council.

Hauptman, P.C., R. Le Blanc and M.B. Wesche (eds.). 1985. Second Language Performance Testing. Ottawa: University of Ottawa Press.

Lo Bianco, Joseph. 1987. National Policy on Languages. Canberra: Australian Government Publishing Service.